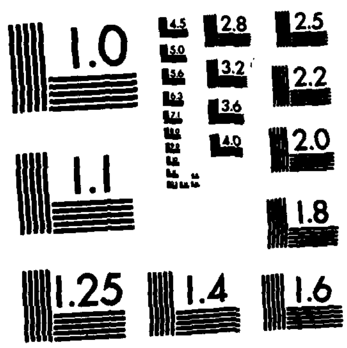
AD-A124 638 THE RESTRICTED STACKELBERG PROBLEM(U) ILLINOIS UNIV AT URBANA DECISION AND CONTROL LAB J T WEN JUN 81 DC-46 N00014-79-C-0424 1/1

UNCLASSIFIED F/G 12/1 NL

END
FILMED
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1

(12)

REPORT DC-46 JUNE, 1981

COORDINATED SCIENCE LABORATORY

DECISION AND CONTROL LABORATORY

THE RESTRICTED STACKELBERG PROBLEM

JOHN TING-YUNG WEN

APPROVED FOR PUBLIC RELEASE. DISTRIBUTION UNLIMITED.

REPORT R-911 UILU-ENG 81-~~2402~~ 2242

UNIVERSITY OF ILLINOIS - URBANA, ILLINOIS

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A124638	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE RESTRICTED STACKELBERG PROBLEM		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER R-911 (DC-46) :UILU-ENG-81-2242
7. AUTHOR(s) John Ting-Yung Wen		8. CONTRACT OR GRANT NUMBER(s) NSF ECS-79-19396 AFOSR-78-3633 N00014-79-C-0424
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Joint Services Electronics Program		12. REPORT DATE June 1981
		13. NUMBER OF PAGES 60
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Stackelberg games
Leader-follower games
Dynamic-to-static
Conversion, team-problem

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Stackelberg equilibrium strategy concept has wide applications in modeling of socio-economic and large scale systems. Its analytic difficulty, however, poses a main drawback. The Restricted Stackelberg Problem (RSP) considers a subset of the Stackelberg strategy where the leader achieves his team cost. It is more analytically tractable and offers a viable alternative.

This report investigates discrete, LQ RSP using dynamic-to-static conversion. The approach reduces the dynamic progression of variables into an augmented

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (continued)

static domain. The static results can then be applied. New sufficient conditions are obtained in simple forms as linear matrix equations for the centralized, decentralized and stochastic centralized information structures. These conditions also encompass the similar work done previously.

Large threat strategy is examined as a possible near-optimal solution. It is found that as the threat tends to infinity, a nonzero offset from the team cost exists for the linear strategy representation, and team cost can be achieved for the discontinuous strategy.

UILU-ENG-81-2242

THE RESTRICTED STACKELBERG PROBLEM

by

John Ting-Yung Wen

This work was supported in part by the Nationsl Science Foundation under Grant ECS-79-19396, in part by the U. S. Air Force under Grant AFOSR-78-3633, and in part by the Joint Services Electronics Program under Contract N00014-79-C-0424.

THE RESTRICTED STACKELBERG PROBLEM

BY

JOHN TING-YUNG WEN

B.Eng., McGill University, 1979

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1981

Thesis Advisor: Professor W. R. Perkins

Urbana, Illinois

ACKNOWLEDGMENT

The author wishes to thank his advisor, Professor W. R. Perkins, for his guidance during the course of this research. He would also like to express his sincere gratitude to Professor J. B. Cruz, Jr. and Professor T. Basar for the generous sharing of their valuable time and ideas.

Special thanks are given to Ms. Rose Harris, Ms. Susan Osmundson, and Ms. Mary Foster for their enormous help in typing and preparing this thesis.

TABLE OF CONTENTS

1. INTRODUCTION

Many problems in control system design or economic system modelling naturally arise in the multiple decision-makers framework. The study of this type of problems is called game theory. The decision-makers are considered as players striving to optimize their respective performance indices under some a priori determined ground rules.

Various types of rules (called strategies) have been developed. Some have only a single performance index; for example, team problem (players optimize the same index but possible under different information) [1], and 2-person zero-sum game (the performance index is the cost of one player and the payoff of the other) [2]. Some games have multiple criteria; for example, the 2-person nonzero -sum game under the Nash equilibrium concept (the players optimize their respective performance simultaneously) [3], or under the Stackelberg equilibrium concept (the leading player optimizes his performance index knowing how the passive player will react) [4].

Game theory, though can be considered as a generalization of the single person, single-criterion control theory, is a great deal more complex. In particular, for the dynamic Stackelberg Game, even in the seemingly simple case of linear-quadratic problem, it is extremely difficult to obtain any analytic solution. Therefore, a modified scheme, the Restricted Stackelberg Problem (RSP) ([5], [6]), is proposed. This is a Stackelberg game with a specific information structure which allows the leader to announce his strategy first but to act only after the follower has acted. By choosing different representations of a given strategy, the leader can manipulate the follower in various ways. In particular, the leader may be able to force the follower to act as if he is also mimimizing the leader's cost. In RSP, we also restrict attention only to those Stackelberg solutions which attain the lower bound of the leader's cost (the team cost). The focus of this report is on RSP for a special class of problem, namely, discrete-time, finite-horizon, linear-quadratic-gaussian.

RSP, if solvable, is a powerful modelling tool. It can be readily applied to many economic and control problems where the hierarchy of operation clearly exists or is desired, and it is analytically tractable. In economics, the government - industry - consumers hierarchy can be naturally posed as a tri-level RSP. The government announces its regulation policy first, the industry then stipulates a pricing strategy based on the announced regulation. The consumers act first by making certain amount of purchase from the industry based on the price of the product or service the industry supplies. In engineering, any large scale system wherein a single centralized controller is impractical can be potentially modelled in the RSP framework with layers of decentralized controllers with different priority of operation.

The investigation in this report is carried out using the dynamic-to-static conversion, which collapses the dynamic evolution of a variable (over finite horizon) into a single vector. A dynamic problem can then be converted into the static domain, and the results proven on this domain can be transferred back to the dynamic domain. One feature of this technique is that it bypasses a great deal of algebra to make the qualitative features more apparent, which is versatile in establishing the existence of solutions.

However, in doing so, it sacrifices the recursiveness of the solutions, which may be a crucial requirement in the implementation of the solutions.

Three classes of information structures are considered: the deterministic centralized, the deterministic decentralized, and the stochastic. Most of the results are obtained for the deterministic centralized information structure. Sufficient conditions for existence of RSP solutions are derived. Some qualitative aspects of RSP are also examined: the dependence of solvability of RSP on the specific choice of information and representation, the stationarity and the convexity conditions, the advantage of linear solutions, and some interpretation of the given conditions. The decentralized problem is approached in the same manner as the centralized case. The results are similar if the initial data distribution is assumed known. The stochastic RSP with perfect state information cannot be solved because of the inability of the leader to detect whether the team solution is enforced or not. To bypass this difficulty, we include both the state and the follower's control to leader's information. The problem then becomes similar to the other cases.

In the situation where the conditions mentioned above are not satisfied, the possibility of the leader using a large threat (penalty to follower's deviation from the team trajectory) strategy to achieve his near-team cost is considered. It is shown that under certain mild conditions, the infinite threat can achieve the team cost for the leader. It is, therefore, reasonable to ask the questions under what conditions can the leader achieve a cost arbitrarily close to his team cost using large but finite threat? It is shown that in general the leader does not possess such a strong position and the case in which it holds is a variety in the parameter space.

This report is structured into four sections. The definitions and problem formulation are stated in Chapter 2. Chapter 3, the main bulk of the work, is devoted to the various cases of RSP. The concluding section, Chapter 4, summarizes the report and points out some future directions.

2. PROBLEM FORMULATION

2.1. Introduction

By an abuse of language, we shall also let RSP stand for the equilibrium strategy to be investigated in this report, which is a restricted version of the Stackelberg equilibrium strategy as briefly discussed in section 1. Stackelberg strategy was introduced by von Stackelberg [12] in the static setting. Generalization to dynamic case was first done in [4]. The idea is that the commanding player (leader) announces his strategy at each stage knowing how the follower will react to his strategy. The follower then optimizes his performance index based on the leader's strategy. This equilibrium strategy concept, although very appealing in terms of modelling, is difficult to solve analytically in general (in the closed loop dynamic case). The difficulty lies in the fact that the principle of optimality fails to apply due to the dependence of the closed loop strategy on the length of the horizon. To circumvent this difficulty a restricted type of Stackelberg strategy is considered in [5], [6]. This strategy concept, RSP, focuses on the Stackelberg pair that achieves the team cost for the leader. The leader, using the non-unique representation of his team strategy, adds on redundant terms that have values zero on the team trajectory. By choosing the appropriate redundancy (or the threat to the follower) the leader may be able to force the follower to act as if he is also optimizing the leader's performance.

In this chapter, we state the general definitions of Stackelberg, Team, and Restricted Stackelberg problems. Then we examine some of the past highlights and show how the present work fits into the lines of development.

2.2. Definitions

Assume some underlying probability space (Ω, F, P) is given. Let $X(0): \Omega \to R^n$, $w(k): \Omega \to R^n$, $V_i(k): \Omega \to R^{y_i}$, $k \in \{0, 1, \text{--}, N-1\}$, $i \in \{1,2\}$, be random variables with respect to (Ω, F, P), whose statics are assumed perfectly known.

We consider a discrete, time-varying, N-stage dynamic system with $u_1(k)$ and $u_2(k)$ as input commands and $W(k)$ as noise disturbance into stage k:

$$x(k+1) = f(k, x(k), u_1(k), u_2(k), W(k)) \quad k = 0, \text{--}, N-1 \tag{2.1}$$

At stage k, assume information vectors $Z_i(k)$ are given:

$$\begin{aligned} Z_i(k) = Z_i(k, x(0), \text{--}, x(k), u_1(0), \text{--}, u_1(k-1), \\ u_2(0), \text{--}, u_2(k-1), y_i(k)) \end{aligned} \tag{2.2}$$

Let $F_i(k)$ be the $Z_i(k)$-generated σ-algebra.

We require that $u_i(k) \in U_i(k)$, where $U_i(k) \triangleq \{\gamma_k^i \mid \gamma_k^i : \Omega \to R^{m_i}, \gamma_k^i(Z_i(k)) \text{ is } F_i(k)\text{-measurable}\}$

The control objective of player i is to find a sequence of admissible controls according to some equilibrium solution concept based on the cost index.

$$\begin{aligned} J_i(\{u_j(k)\}_{k=0,\ j=1}^{N-1,\ 2}) = E\{\sum_{k=0}^{N-1} L_{i,k}(x(k), u_1(k), \\ u_2(k), k) + \varphi_i(x(N))\} \end{aligned} \tag{2.3}$$

We now define the following equilibrium solution concepts.

Definition 2.1

$\{u_1(k), u_2(k)\}_{k=0}^{N-1}$ is a closed loop Stackelberg sequence with player 1 as leader, player 2 as follower if it solves

$$\underset{\substack{u_1(k)\in U_1(k)\\ k=0,1,\text{--},N-1}}{\text{Min}} J_1\left(\{u_1(k)\}_{k=0}^{N-1}, \{\hat{u}_2(k, \{u_1(k)\}_{k=0}^{N-1})\}_{k=0}^{N-1}\right)$$

where $\{\hat{u}_2(k, \{u_1(k)\}_{k=0}^{N-1})\}_{k=0}^{N-1}$ is

$$\{\hat{u}_2(k, \{u_1(k)\}_{k=0}^{N-1}) = \underset{u_2(k)\in U_2(k)}{\text{arg Min}} J_2^2(\{u_1(k)\}_{k=0}^{N-1}, \{u_2(k)\}_{k=0}^{N-1})$$

Definition 2.2

$\{u_1^t(k), u_2^t(k)\}_{k=0}^{N-1}$ is a team solution pair for the leader if it solves

$$\underset{\substack{u_i(k)\in U_i(k)\\ i=1,2\\ k=0,\text{--},N-1}}{\text{Min}} J_1(u_1(0), \text{--}, u_1(N-1), u_2(0), \text{--}, u_2(N-1))$$

Definition 2.3

Let $Z_1(k, u_2(k), u_2(k-1), \text{--}, u_2(0))$ be some information set.

Then, $\{u_1(Z_1(k, u_2(k), \text{--}, u_2(0)), u_2(k)\}_{k=0}^{N-1}$ is the solution of RSP if

it solves

$$\{u_2^t(k)\}_{k=0}^{N-1} = \underset{\substack{u_2(k)\in U_2(k)\\ k=0,\text{--},N-1}}{\text{arg min}} J_2(u_1(Z_1(0, u_2(0)) \text{--}, u_1(Z_1(N-1),$$

$$u_2(N-1), \text{--}, u_2(0))), u_2(0), \text{--}, u_2(N-1))$$

and $u_1^t(k) = u_1(Z(k, u_2^t(k), u_2^t, (k-1), \text{--} u_2^t(0)))\quad \forall k\in\{0,1, \text{--}, N-1\}$

2.3. Problem Formulation

In this report, we consider specifically the discrete, finite-horizon, linear-quadratic deterministic and stochastic gaussian systems. The technique employed is the dynamic-static conversion. The time evolution is collapsed into a single long column vector. The system can then be viewed as a static entity; however, the relationship between these time-vectors has to be restricted by causality. Thus, the techniques available in the static case can be readily applied under the causality constraint. The system under consideration is described by

$$X(k+1) = A(k)X(k) + B_1(k)\, U_1\,(k) + B_2(k)\, U_2(k) + W(k) \qquad k=0,\ 1,\ --,\ N-1 \tag{2.4}$$

$U_1\,(k)$, $U_2\,(k)$ are the controls of players 1 and 2 respectively at stage k. The cost function of player i is given as:

$$J_i\ (\{U_i\ (k)\},\ \{U_j\ (k)\}) = E\ \{X'(N)\ Q_i(N)X(N) + \sum_{k=0}^{N-1}[X'(k)\ Q_i(k)X(k) + U_i'\ (k)\ R_{ii}(k)U_i(k) + U_j'\ (k)R_{ij}(k)U_j(k)]\},\quad i,j=1,2,\ i=j. \tag{2.5}$$

Assume also

$Q_i\ (k) \geq 0 \qquad R_{ij}\ (k) > 0 \qquad i,\ j = 1,2$

$W(k) \sim N\ (0,\ \Sigma_w(k))$

$X_o \sim N\ (\bar{X}_o,\ \Sigma_o)$

Note that in the usual Nash formulation, R_{12} need not be positive definite. It will be shown in Chapter 3 that this is a necessary condition for RSP.

We convert the above dynamic system into the static domain via the following:

Define

$$X = \begin{bmatrix} X(0) \\ \vdots \\ X(N) \end{bmatrix} \quad U_i = \begin{bmatrix} U_i(0) \\ \vdots \\ U_i(N-1) \end{bmatrix} \quad W = \begin{bmatrix} W(0) \\ \hline \vdots \\ \hline \sum_{j=0}^{N-1} \Phi(N-1,j)W(j) \end{bmatrix} \tag{2.6}$$

Then the state equation collapses to

$$X = \begin{bmatrix} I \\ \hline \vdots \\ \hline \Phi(N,0) \end{bmatrix} X(0) + \sum_{i=1}^{2} \begin{bmatrix} 0 & & 0 \\ B_i(0) & \ddots & \\ \vdots & & 0 \\ \Phi(N-1,0)B_i(0) & \cdots & B_i(N-1) \end{bmatrix} U_i + \begin{bmatrix} W(0) \\ \vdots \\ \hline \sum_{j=0}^{N-1} \Phi(N-1,j)W(j) \end{bmatrix}$$

$$\stackrel{\Delta}{=} D\,X(0) + \sum_{i=1}^{2} H_i U_i + W \tag{2.7}$$

Similarly,

$$J_i = E\,[\,X' Q_i X + U_i' R_{ii} U_i + U_j' R_{ij} U_j\,] \tag{2.8}$$

$$Q_i = \text{diag}\ [\,Q_i(0), --, Q_i(N)\,]$$

$$R_{ij} = \text{diag}\ [\,R_{ij}(0), --, R_{ij}(N-1)]$$

Causality is an important property of the functional mappings in this setting. It is characterized in a simple way for matrices, namely, the block lower triangularity implies causality.

We therefore define the following:

Definition 2.5:

A matrix $F = [f_{ij}]$, f_{ij} = some matrix with known dimension

① causal if $f_{ij} = 0 \quad \forall j > i$

② strictly causal if $f_{ij} = 0 \quad \forall j \geq i$

The advantage of working in this pseudo-static domain, as stated before, is the simplicity of algebra and the applicability of static result in a straight-forward manner. However, same results may also be obtained by using dynamic programming.

2.4. Past Development in RSP and Team Problem

To study RSP, it is certainly necessary to solve the corresponding team problem. For centralized, deterministic information structure, the team problem is the same as the optimal control problem, the solution of which is of course well known. Unfortunately, in the general decentralization case, it requires an infinite dimensional filter to generate all the estimates. Therefore, due to realizability, additional assumptions on the information have to be made. One type of assumption ([18], [19]) restricts information to that generated by a finite-dimensional, linear filter, the optimal solution can then be found. Another type of assumption ([8], [9], [10]) is the nested information where observations are shared with one-step delay. This report uses the similar idea as ([18], [19]). Parameter optimization is used to find the best linear strategies. Due to the conversion to the static domain, the sufficient conditions are stated in particularly simple forms.

RSP is formally investigated by Basar [5] and Papavassiloupous [6,7] under perfect state information. Sufficient conditions are obtained in each case for the RSP solution to exist. However, some issues are left mostly unaddressed: the effect of leader's information structure on the solvability of RSP, RSP under large threat, possibility of suboptimal RSP solution should the sufficiency conditions fail, the qualitative interpretation of the conditions etc. The stochastic RSP given the state information only is in general unsolved and appears unsolvable in the dynamic case. It is solvable in the static setting, however, as in [13], [14]. In this report, we include the follower's past control in the leader's information structure and are, therefore, able to solve the problem. We also solve the deterministic decentralized RSP under the linear strategy constraint (to bypass the difficulty in the general team problem). The centralized deterministic RSP is also studied, and some of the previously little touched issues are explored. However, there still exists a great deal of open problems, especially with regard to the near-optimal solutions in the stochastic RSP.

3. RESTRICTED STACKELBERG PROBLEM

3.1. Introduction

Dynamic RSP has been studied in [5], [6], in which, conditions for enforcing the team solution for the leader are obtained under perfect state information in the deterministic problem. Some results on the static stochastic RSP are presented in [13], [14]. Here we first examine the deterministic RSP under various information structures and then the stochastic RSP under a specific information pattern.

The RSP is approached as follows:

1. Solve the team problem for the leader under the given information structure.
2. Choose one representation of leader's team strategy such that it is dependent on the follower's decision non-trivially.
3. Find conditions this representation must satisfy such that follower's decision from his own optimization coincides with the team solution.

We shall consider the following information structures:

(Let $Z_i(k) \triangleq$ information available to $U_i(k)$)

Deterministic

a. $Z_1(k) = [U_2(k), U_2(k-1), --, U_2(0), X_o]$

$Z_2(k) = [X(k), X(k-1), --, X_o]$

b. $Z_1(k) = [X(k), X(k-1), --, X_o]$

$Z_2(k) = [X(k), X(k-1), --, X_o]$

c. $Z_1(k) = [U_2(k), U_2(k-1), --, U_2(0), X(k), X(k-1), --, X_o]$

$Z_2(k) = [X(k), X(k-1), --, X_o]$

d. $Z_1(k) = [Y_1(k), Y_1(k-1), --, Y_1(0)]$

$Z_2(k) = [Y_2(k), Y_2(k-1), --, Y_2(0)]$

($Y_1(\cdot)$ $Y_2(\cdot)$ are non-nested.)

Stochastic

e. $Z_1(k) = [U_2(k), --, U_2(0), X(k), X(k-1), --, X(0)]$

$Z_2(k) = [X(k), X(k-1), --, X(0)]$

Note:

1. We have allowed $U_1(k)$ to be dependent on $U_2(k)$. This certainly is not physically possible since $U_1(k)$ needs a nonzero amount of time for computation. However, here we assume that the interval between two stages is long relative to the delay, thus, we can consider the strategies $U_1(k)$ and $U_2(k)$ as being implemented at the same stage. If precision is needed to include this delay in the model, we can subdivide the interval and let $U_1(k)$ depend on $U_2(k-1)$,--, $U_2(0)$ only. In either case, the subsequent results are the same. Care only needs to be taken to restrict the matrix coefficient mapping U_2 to U_1 to be causal (block lower triangular) in the former case and strictly causal (strictly block lower triangular) in the latter.
2. Cases (a), (b), (c) are considered to examine the impact of leader's information structure on the solvability of RSP. Case (d) in the general deterministic decentralized information, in which, the team solution cannot be obtained in general. Therefore, the best linear

strategies are derived and RSP is solved based on the assumption that leader enforces these strategies. In the stochastic RSP with only the state information, the leader has no way of enforcing his team solution since the team trajectory depends on the sample path of a gaussian random process. In (e), we include the follower's past strategies as well so that the leader can use them for the threat.

3. The solvability of RSP is also viewed from the asymptotic behavior of the follower's strategy as a function of the strength of leader's threat. It is shown that under some mild conditions, if the leader threatens to play an infinite control for any deviation of the follower's strategy from the desired value, leader team solution can be enforced. Since infinite gain is not physically possible, we examine the possibility of a large, finite threat. It is shown, with aid of an example, that arbitrary closeness to the team cost may not be forced with a linear representation no matter how large (but finite) the threat is. However, if discontinuous strategies are allowed for the leader, it can be shown that arbitrary closeness to the team cost can then be achieved with a large, finite threat.

3.2. Centralized Deterministic RSP

RSP under information structures (a), (b), (c) is examined in this section. Non-void conditions are expected to exist in each case since the expected team values (of states and follower's control) are known exactly, and an impulsive punishment can be used to threaten the follower. Suppose finite gain is required for the leader, then the leader's strategy has to satisfy some conditions.

We first derive the team solution for the leader. Then with the general structure on the leader's strategy (only differentiability is assured), sufficient conditions (first order stationarity and second order comoxity conditions) are derived for the existence of finite solutions.

The team solution can be easily obtained using dynamic programming, but to stay consistently with the converted scheme, we shall derive it under the present setting.

3.2.1. Team solution

In this section, we derive the open loop and closed loop team solution for the leader using the converted system.

Consider (2.7), (2.8) as a static team optimization problem with cost.

$$J_1 = [(DX_o + H_1 U_1 + H_2 U_2)' Q_1 (DX_o + H_1 U_1 + H_2 U_2) + U_1' R_{11} U_1 + U_2' R_{12} U_2] \qquad (3.1)$$

Using the deterministic counterpart of Radner's Theorem [1], set

$$P\,[\nabla_{U_i} J_i] = 0 \qquad i = 1,2 \tag{3.2}$$

where

$$P = \text{diag}\,[P(0),--,P(N-1)] \tag{3.3}$$

$$P(k) = \text{projection onto the space spanned by } [X(0),--,X(k)] \tag{3.4}$$

Then,

$$\begin{aligned} U_i^{\,t} &= -P\,[(H_i{}' Q_i H_i + R_{ii})^{-1} H_i{}' Q_i (DX_o + H_j U_j)] \\ &= -P\,[R_{ii}{}^{-1} H_i{}' Q_i X] \end{aligned} \tag{3.5}$$

Note that we have used the assumption $R_{12} > 0$, since otherwise impulsive $U_2^{\,t}$ may result.

We notice that $U_i^{\,t}$ has a non-causal dependence on X. Therefore, we use the following transformation to obtain a causal representation.

<u>Proposition 3.1</u>

Given U_i as

$$U_i = -P \; R_{ii}{}^{-1} H_i{}' Q_i X$$

Assume

$$G_k = \begin{bmatrix} I - d^{(1)}_{k,k+1} B_1(k) & -d^{(1)}_{k,k+1} B_2(k) \\ -d^{(2)}_{k,k+1} B_1(k) & I - d^{(2)}_{k,k+1} B_2(k) \end{bmatrix} \text{ is invertible} \tag{3.6}$$

where

$$d^{(i)}_{k,k+1} = \sum_{\ell=k+1}^{N} [R_{ii}{}^{-1}(k)\, B_i{}'(k)\, \Phi'(\ell-1,k)\, Q_i(\ell) \prod_{j=k+1}^{\ell-1} (A(j) + B_1(j)\, g_1(j) + B_2(j)\, g_2(j))] \tag{3.7}$$

$$\begin{bmatrix} g_1 \ (j) \\ g_2 \ (j) \end{bmatrix} = G_j^{-1} \begin{bmatrix} d_{j,j+1}^{(1)} \\ d_{j,j+1}^{(2)} \end{bmatrix} \quad A(j) \tag{3.8}$$

Then $U_i\ (k) = g_i\ (k)\ X\ (k)$ is a causal version of (3.5).

Proof

See Appendix 1. Q.E.D.

We write the closed loop solution as

$$U_i^t = G_i X \tag{3.9}$$

where G_i is block diagonal with components as calculated in Proposition (3.1).

It is well known that the open loop and closed loop versions of the control lead to the same state trajectory. For the open loop:

$$X^t = (I - H_1 G_1 - H_2 G_2)^{-1} D X_o \tag{3.10}$$

$$\begin{aligned} U_i^t &= G_i X^t \\ &= G_i (I - H_1 G_1 - H_2 G_2)^{-1} D X_o \\ &\stackrel{\Delta}{=} G_i^0 X_o \end{aligned} \tag{3.11}$$

Remarks

With state feedback the first order condition should actually be

$$(H_i{}' Q_1 H_i R_{ii}) U_i + P \Big\{ [H_i{}' Q_1 (DX_o + H_j U_j)] + \nabla_{u_i} U_j [(H_j{}' Q_1 H_j + R_{ij}) U_j + H_j{}' Q_1 (DX_o + H_i U_i)] \Big\} = 0$$

One solution is the pair (3.5), which is the open loop solution. The only case non-uniqueness may occur is when $\nabla_{u_i} U_j \nabla_{u_j} U_i = I$. But

$$\nabla_{u_i} U_j \nabla_{u_j} U_i = \nabla_x U_j H_i \nabla_x U_j H_j$$

and $\text{Ker} \quad H_j \neq 0$

$$\therefore \nabla_{ui} U_j \nabla_{uj} U_i \neq I$$

The closed loop team solution is therefore the same as the open loop solution in the sense they both achieve the lower bound of leader's cost. It is, however, immediately noticed that such advantage is not enjoyed in the multicriteria case, e.g., Nash or Stackelberg. In these cases, nested information is used to eliminate $\nabla_{u_i} U_j$ terms.

3.2.2. Conditions for enforcing the team solution

(1) Sufficient conditions

In this section, we derive the sufficient conditions for the leader to enforce his team solution using non-uniqueness of representation of his team strategy. The condition will be composed of the first order stationarity condition, the second order convexity condition, and the additional assumption that if the leader's strategy is fixed, follower's optimization admits his part of the team solution as a globally mimimizing solution. The stationarity and the convexity conditions are investigated further for any differentiable representation of leader's team strategy. More specific conditions are then obtained for each case. For linear representation, it is shown that the convexity condition and the global minimum condition are always satisfied.

We choose a representation of U_1 as

$$U_1 = Y_1 (Z_1 (U_2)) \tag{3.12}$$

where $Y_1 (\cdot)$ is chosen to satisfy

$$(1) \quad \arg\min_{U_2} J_2 (Y_1 (Z_1 (U_2)), U_2) = U_2^t \tag{3.13}$$

$$(2) \quad Y_1 (Z_1 (U_2^t)) = U_1^t \tag{3.14}$$

We define functions with property (2) as class-T functions. The objective here is to find sufficient conditions for φ under information structures (a), (b), (c), given $Y_1 (\cdot)$ a class-T function. The information available to the leader, $Z_1 (\cdot)$ is some function dependent on U_2 in a causal manner. If Z_1 is independent of U_2, leader will have no way of influencing the follower's optimization.

We now state the sufficient conditions and the proof:

<u>Theorem 3.2</u>

Assume

(i) $Y (Z_1)$ is a causal, differentiable, class-T function

(ii) $J_2 (Y (Z_1), U_2)$ is convex

(iii) $Z_2 \supset \{X_o\}$

(iv) $\exists \quad F = \nabla_{U_2} Y(Z_1) \Big|_{U_2 = U_2^t}$ $\quad (\nabla_{U_2} \triangleq$ total differential with respect to $U_2)$ (3.15)

$$\ni [(R_{22}G_2 + H_2'Q_2) + F^* (R_{21}G_1 + H_1^* Q_2)] (I-H_1G_1-H_2G_2)^{-1} D=0 \tag{3.16}$$

Then $U_1 = Y(Z_1)$ will force U_2 to adopt U_2^t.

Proof:

If $U_1 = \gamma\ (Z_1)$ is a class-T function, J_2 is convex, $P\ [\nabla_{U_2}\ J_2] \Big|_{U_2 = U_2^t} = 0,$

then the global mimimum of $J_2\ (\gamma\ (Z_1)\ U_2)$ is attained with the pair $(U_1^t,\ U_2^t)$. Therefore, it sufficies to show that $P_{z_2}\ [\nabla_{U_2}\ J_2] \Big|_{U_2 = U_2^t} = 0$

implies condition (iv). (P_{z_2} is the projection onto the space spanned by z_2.)

We know that knowing X_o is sufficient to achieve the lower bound of the cost function in a deterministic control problem. And since $\{X_o\}$ and $\{X\}$ are equivalent in the sense that they both achieve the minimum, we can substitute P (projection as in (3.3), (3.4)) for P_{z_2}.

$$P\ [\nabla_{U_2}\ J_2] = P\ [(\nabla_{U_2} X)'\ Q_2 X + (\nabla_{U_2} Y)'\ R_{21} Y + R_{22} U_2] = 0 \qquad (3.18)$$

$$\nabla_{U_2} X = H_2 + H_1\ \nabla_{U_2}\ Y$$

$$\therefore\ P\ [(\nabla_{U_2}\ \gamma\ (Z_1))'\ (H_1' Q_2 X + R_{21}\ \gamma(Z_1)) + H_2' Q_2 X + R_{22} U_2] = 0$$

Let $U_2 = G_2 X^t$, then $\gamma\ (U_2) = G_1 X^t$

It is sufficient then

$$[(\nabla_{U_2}\ \gamma(Z_1)|)'\ (H_1' Q_2 + R_{21}\ G_1) + (H_2' Q_2 + R_{22}\ G_2)]\ X^t = 0$$

$$U_2 = U_2^t$$

$$X = X^t$$

for all possible X^t.

$$X^t = (I-H_1G_1 - H_2G_2)^{-1} D X_o$$

and there is no restriction on X_o.

$$\therefore [(\nabla_{U_2} \gamma (Z_1)|)' (H_1'Q_2 + R_{21}G_1) + (H_2'Q_2 + R_{22}G_2)] (I-H_1G_1-H_2G_2)^{-1} D = 0$$

$$U_2 = U_2^t \qquad\qquad \text{Q.E.D.}$$
$$X = X^t$$

<u>Discussion</u>:

1. The above theorem holds for the information structures (a), (b), (c). However, the solvability differs on each case due to the different $\nabla_{U_2} \gamma (Z_1)$ expressions. Let F be defined as in (3.16).

For (a), $Z_1 = \{U_2\ X_o\}$, $Z_2 = \{X_o\}$:

$$\nabla_{U_2} \gamma (U_2) \Big|_{U_2 = U_2^t} = F \tag{3.19}$$

For (b), $Z_1 = \{X\}$, $Z_2 = \{X\}$

$$\nabla_{U_2} \gamma (X) = \nabla_X \gamma (X) \ \nabla_{U_2} X$$

$$\nabla_{U_2} X = H_1 \nabla_X \gamma (X) \nabla_{U_2} X + H_2$$

$$= (I-H_1\nabla_X \gamma (X))^{-1} H_2$$

We need

$$\nabla_X \gamma (X) \Big|_{X = X^t} (I-H_1\nabla_X \gamma (X)\Big|_{X = X^t})^{-1} H_2 = F$$

or

$$(\nabla_X \gamma (X)) \Big|_{X = X^t} (H_2 + H_1 F) = F \tag{3.20}$$

For (c), $Z_1 = \{U_2, X\}$, $Z_2 = \{X\}$

Let $\nabla^{(i)}$ denote gradient with respect to ith variable, ∇ denotes the total gradient.

Then

$$\nabla_{U_2} X = H_1 \left(\nabla_X^{(1)} \gamma(X, U_2)\, \nabla_{U_2} X + \nabla_{U_2}^{(2)} \gamma(X,U_2)\right) + H_2$$

$$= \left(I - H_1 \nabla_X^{(1)} \gamma(X, U_2)\right)^{-1} \left(H_1 \nabla_{U_2}^{(2)} \gamma(X,U_2) + H_2\right)$$

$$\nabla_{U_2} \gamma(X,U_2) = \nabla_X^{(1)} \gamma(X,U_2)\, \nabla_{U_2}^{(2)} \gamma(X,U_2)$$

$$= \nabla_X^{(1)} \gamma(X,U_2) \left(I - H_1 \nabla_X^{(1)} \gamma(X,U_2)\right)^{-1}$$

$$\left(H_1 \nabla_{U_2}^{(2)} \gamma(X,U_2) + H_2\right) + \nabla_{U_2}^{(2)} \gamma(X,U_2)$$

$$\therefore \left.\nabla_X^{(1)} \gamma(X,U_2)\right|_{\substack{X = X^t \\ U_2 = U_2^t}} (H_2 + H_1 F) + \left.\nabla_{U_2}^{(2)} \gamma(X,U_2)\right|_{\substack{X = X^t \\ U_2 = U_2^t}} = F \qquad (3.21)$$

Note that (3.21) reduces to (3.19) or (3.20) if $\nabla_X^{(1)} \gamma(X,U_2)$ or $\nabla_{U_2}^{(2)} \gamma(X,U_2)$ is set to zero respectively.

As a design method, F is first solved from (3.16). Then depending on the information structure, the appropriate $\gamma(\cdot)$ can be chosen. However, the ability to choose $\gamma(\cdot)$ differs in each case. Given F: in (a), (c), $\nabla\gamma$ can always be solved.

in (b), it is necessary and sufficient Ker $(H_2 + H_1 F) \subset$ Ker F. (3.22)

Even though, given F, (c) does not seem to offer anything extra in terms of the solvability of RSP, the additional term in (3.21) does provide freedom to possibly attain other desirable features (e.g., sensitivity, convexity, etc. In comparing the three information structures, we conclude that (b) is more restrictive than (a) and (c). (c) offers additional freedom to fine tune other features of the solution.

2. Some of the assumptions in the theorem may seem restrictive, however, in fact, they are due to reasonable necessity.
   - Assumption (i) restructs the class of leader's strategies, causality, and class-T are necessary, differentiability helps to carry out optimization analytically. In general, these are not very stringent since a large class of functions still remain.
   - Assumption (ii) states the convexity condition. It is necessary to guarantee the existence of at least one local relative minimum.
   - Assumption (iii) restricts the theorem to the perfect state information. The decentralized case will be treated separately later.
   - Assumption (iv) is the stationarity condition. The expression in (3.16) is necessary and sufficient provided F is finite (the infinite gain case will be discussed later).
3. Note that $\nabla_U \, ; \, (Z_1) \Big| =$ constant matrix imposes a strong restriction

$$X = X^t$$

$$U = U^t$$

on non-linear functions, since it implies that all the terms with order higher than or equal to two will have to vanish on the team trajectory. Furthermore, it will be shown that convexity condition also becomes very restrictive for the non-linear functions. Both of these suggest that linear strategy as the ideal condidate since they are trivially satisfied. However, for the state information case, it is shown in [21] that, in certain examples, only non-linear solutions exist. Another property to notice is that if an F exists in (3.16), it is causal. This is certainly necessary for a linear strategy.

(ii) Stationarity and Convexity:

In this section, we examine the stationarity and the convexity conditions ((ii), (iv) respectively in more detail. The stationarity condition is expressed in geometric language. The sufficient condition for convexity is derived.

- Stationarity:

In (3.16), we have $\frac{N(N+1)}{2}$ $(m_1 \times m_2)$ unknowns (due to the causal structure of F), and N $(n \times m_2)$ equations. Assume that all equations are independent. Then we require $N \geq \frac{2n}{m_1} - 1$. If equality holds, the solution F is unique.

If strict inequality holds, there are, in general, infinitely many solutions. The advantage of this freedom and the ways of utilizing it requires further study. If inequality fails, we then have to solve F as a function of X_o, in which case, $N \geq \frac{2}{m_1} - 1$ always holds.

In the case of state information (information structure (b)), we have the additional equation (3.20) to solve. Observe that in the case of linear strategy, $\nabla_x(\gamma(x))$, a constant matrix must have the last n columns equal to zero due to causality restriction. Therefore, for (3.20), the last m_2 columns of F (Nm_1m_2 elements) must also be zero. We have then

$$\frac{N(N+1)}{2}\, m_1m_2 - Nm_1m_2 = \frac{N(N-1)}{2}\, m_1m_2$$

unknowns and Nnm_2 equations. Thus, we need $N \geq \frac{2n}{m_1} + 1$ in general, i.e. if all equations are independent. Putting together the above constraint and (3.16), we have:

Proposition 3.3

The equation (3.16) has a solution F if

(i) $\ker(R_{21}G_1+H_1'Q_2) \cap \mathrm{Im}(I-H_1G_1-H_2G_2)^{-1}D \subset \ker(R_{22}G_2+H_2'Q_2) \cap \mathrm{Im}(I-H_1G_1-H_2G_2)^{-1}D$ (3.23)

(ii) $\mathrm{rank}(R_{21}G_1+H_1'Q_2)(I-H_1G_1-H_2G_2)^{-1}D \geq \mathrm{rank}(R_{22}G_2+H_2'Q_2)(I-H_1G_1-H_2G_2)^{-1}D$ (3.24)

(iii) $N \geq \frac{2n}{m_1} - 1.$

Proof: Write (3.16) as

$$A = -F'B.$$

It is necessary and sufficient that

$$\mathrm{Im}A = \mathrm{Im}F'B$$

and

$$\ker A = \ker F'B.$$

The first condition and one direction of the second condition ($\ker A \subseteq \ker F'B$) is taken care of by choosing F appropriately and condition (ii).

We know this is possible because $N \geq \frac{2n}{m_1} - 1$ implies number of unknowns is greater than the number of constraints.

Now we need $\ker A \subseteq \ker F'B$. Since F is free to be chosen, we only need to require $\ker A \subseteq \ker B$. Substitute for A and B with respective expressions, the result follows. Q.E.D.

Condition (i) means that for all possible team trajectories, x^t, $u_1^t R_{21} u_1^t + u_1^t H_1' Q_2 x^t = 0$, $u_2^t R_{22} u_2^t + u_2^t H_2' Q_2 x^t = 0$. This is certainly necessary since if $u_1^t R_{21} u_1^t + u_1^t H_1' Q_2 x^t = 0$ and $R_{22} u_2^t + H_2' Q_2 x^t = \varepsilon \neq 0$, then u_2^t is not the optimal solution for the follower, while $u_2^t = P[R_{22}^{-1}(H_2' Q_2 x^t - \varepsilon)]$ is. Condition (ii) simp simply requires the number of unknowns to be greater than the number of equations.

- Convexity

Recall that U_2 is defined as the set of all functions, u_2, measurable with respect to the σ-algebra generated by the information structure. U_2 is certainly convex since if $u_2^{(1)}, u_2^{(2)} \in U_2$, $\alpha u_2^{(1)} + (1-\alpha) u_2^{(2)} \in U_2$. We have assumed the differentiability of $\gamma(z_1)$, therefore, $J_2(\gamma(z_1), u_2)$ is convex over U_2 if and only if $\nabla^2_{u_2} J_2(\gamma(z_1), u_2) \geq 0$

$$\nabla_{u_2} J_2(\gamma(z_1), u_2) = [(\nabla_{u_2} \gamma(z_1))'(R_{21}\gamma(z_1) + H_1' Q_2 x) + R_{22} u_2 + H_2' Q_2 x]$$

$$\begin{aligned} \nabla^2_{u_2} J_2(\gamma(z_1), u_2) = {} & R_{22} + (H_1 \nabla_{u_2} \gamma(z_1) + H_2)' Q_2 (H_1 \nabla_{u_2} \gamma(z_1) + H_2) \\ & + \frac{1}{2} [(\nabla^2_{u_2} \gamma(z_1))(R_{21}\gamma(z_1) + H_1' Q_2 x) + (R_{21}\gamma(z_1) \\ & + H_1' Q_2 x)' \nabla^2_{u_2} \gamma(z_1)]. \end{aligned} \tag{3.25}$$

Since R_{22}, Q_2, R_{21} are all positive definite or positive semi-definite, we only need

$$(\nabla^2_{u_2}\gamma(z_1))(R_{21}\gamma(z_1)+H_1'Q_2x)+(R_{21}\gamma(z_1)+H_1'Q_2x)'(\nabla^2_{u_2}\gamma(z_1))\geq 0$$

$$\forall x \text{ and } \gamma(z_1) \text{ generated by } u_2\in U_2. \tag{3.26}$$

If $\nabla^2_{u_2}\gamma(z_1)\neq 0$, the above condition is very difficult to be satisfied. The reason is that if we consider $\nabla^2_{u_2}\gamma(z_1)$ evaluated at a particular u_2, we only need $(R_{21}\gamma(z_1)+H_1'Q_2x)|_{u_2}$ to have the opposite sign of $\nabla^2_{u_2}\gamma(z_1)$ in one of its orthogonal coordinates. This immediately suggests the desirability of the linear strategy, since $\nabla^2_{u_2}\gamma(z_1)=0$ in that case

(iii) Linear Strategies

From the discussion in the previous sections, we see that the non-linear representation of leader's strategy does not offer any advantage; in fact, considerable care needs to be taken for convexity. Therefore, we now specialize our attention to linear representation only.

Proposition 3.4

Assume

(i) $\gamma(z_1)$ is a causal, differentiable, class-T function

(ii) $z_2\supset\{x_o\}$

(iii) $\exists F \ni$

$$[(R_{22}G_2+H_2'Q_2)+F'(R_{21}G_1+H_1'Q_2)](D+H_1G_1^o+H_2G_2^o) = 0. \tag{3.27}$$

(iii)' If $z_1=\{x\}$, $z_2=\{x\}$, assume $\exists K\ni$

$$(K+G_1)(H_2+H_1F) = F. \tag{3.28}$$

(iii)" If $z_1=\{x,u_2\}$, $z_2=\{x\}$, assume $\exists K\ni$

$$(G_1-KG_2)(H_2+H_1F)+K = F. \tag{3.29}$$

Then,

$$\text{for } z_1 = \{u_2\},\ z_2 = \{x\}, \quad u_1 = F(u_2 - G_2^o x_o) + G_1^o x_o; \tag{3.30}$$

$$\text{for } z_1 = \{x\},\ z_2 = \{x\}, \quad u_1 = K(x - x^t(x_o)) + G_1 x; \tag{3.31}$$

$$\text{for } z_1 = \{u_2, x\},\ z_2 = \{x\},\ u_1 = K(u_2 - G_2 x) + G_1 x; \tag{3.32}$$

will force u_2 to adopt u_2^t, respectively.

Proof: Substitute the expression of $u_1(\cdot)$ into Theorem 3.2, the result then follows. Q.E.D.

Note that the convexity condition vanishes due to the fact $\nabla^2_{u_2} u_1(z_1) = 0$. The conditions are easy to verify since they only involve linear equations. The gain matrices are all causal (if they exist), therefore, the solution is also realizable (causality is ensured $\forall G_1, G_2$ in diagonal or noncausal representation).

3.2.3. Examples

We shall examine some simple scalar, 2-stage examples. Team and RSP under information structures (a),(b) are solved using the technique derived before. The RSP solutions are verified by substituting them back into J_2 and solve for the optimal u_2. The effect of weighting matrix coefficients on the solvability and the implication of different information structures ((a) vs. (b)) are clearly illustrated.

Consider a scalar, 2-stage system

$$x(2) = 2x(1) + u_1(1) + u_2(1)$$
$$x(1) = x(0) + u_1(0) - u_2(0)$$

$$J_1 = 2x^2(2) + x^2(1) + x^2(0) + 2u_1^2(1) + u_1^2(0) + u_2^2(1) + u_2^2(0)$$

$$J_2 = x^2(2) + x^2(1) + 2x^2(0) + au_1^2(0) + bu_1^2(1) + cu_2^2(0) + du_2^2(1).$$

Apply static-conversion

$$x = \begin{bmatrix} x(0) \\ x(1) \\ x(2) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} x(0) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} u_1(0) \\ u_1(1) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} u_2(0) \\ u_2(1) \end{bmatrix}$$

$$J_1 = x' \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} x + u_1' \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} u_1 + u_2' u_2$$

$$J_2 = x' \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} x + u_1' \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} u_1 + u_2' \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} u_2.$$

<u>Team</u>

The non-causal control laws, from (3.5), are

$$u_1 = -\begin{bmatrix} 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} x$$

$$u_2 = -\begin{bmatrix} 0 & -1 & -4 \\ 0 & 0 & 2 \end{bmatrix} x.$$

Using Proposition 3.1, we transform them to the causal representation

$$u_1 = \begin{bmatrix} -3/7 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \end{bmatrix} x$$

$$u_2 = \begin{bmatrix} 3/7 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} x.$$

The team trajectory is

$$x = \begin{bmatrix} 1 \\ 1/7 \\ 1/14 \end{bmatrix} x(0)$$

and the open loop control law is

$$G_1^o = \begin{bmatrix} -3/7 \\ -1/7 \end{bmatrix}$$

$$G_2^o = \begin{bmatrix} 3/7 \\ -1/7 \end{bmatrix}.$$

<u>RSP</u>

We now assume linear strategy and apply Proposition 3.4.

- Information structure (a)

Let

$$u_1 = F(u_2 - G_2^o x_o) + G_1^o x_o .$$

F satisfies

$$[F'(H_1'Q_2 + R_{21}G_1) + (H_2'Q_2 + R_{22}G_2)](D + H_1G_1^o + H_2G_2^o) = 0$$

Restrict F to the causal structure $F = \begin{bmatrix} f_1 & 0 \\ f_2 & f_3 \end{bmatrix}$. Substitute in numerical values, we obtain

$$(4-6a)f_1 + (1-b)f_2 = 4-6c$$

$$f_3 = \frac{2d-1}{1-b} .$$

We notice immediately that F is nonunique (3 variables and 2 equations), however, given a,b,c,d, f_3 is unique. This points out the possibility that given the weighting parameters, we can tune F to achieve better performance in, say, parameter sensitivity; or, given the desired F, we can tune the parameters. We can also check that $N \geq \frac{2n}{m_1} - 1$ $(2 > 1)$. Therefore, provided equations are all independent, we should have 1 (2-1) degree of freedom).

For some values of a,b,c,d, F may not exist at all (this points to the importance of suboptimal strategy) even in this simple example, e.g., $b=1$, $a=\frac{4}{6}$, $c\neq\frac{4}{6}$. However, we are able to say F exists generically. To verify that the stated strategy does enforce team, we substitute in some numerical values for (a,b,c,d) and solve for the optimal u_2.

- $a=c=d=1 \quad b=0 \qquad F=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ set $f_2=0$ arbitrarily (diagonal structure)

$$u_1(0) = u_2(0) - \frac{6}{7}x(0)$$

$$u_1(1) = u_2(1) + \frac{1}{14}x(0).$$

Substitute in J_2 and set $\frac{\partial J_2}{\partial u_2(0)} = \frac{\partial J_2}{\partial u_2(1)} = 0$, we obtain

$$u_2(0) = \frac{3}{7}x(0)$$

$$u_2(1) = -\frac{x(0)}{7}$$

as expected.

- $a=b=0$ ($R_{21}=0$, i.e., u_1 is not penalized directly in J_2) $c=d=1$.

Set $f_2=0$ arbitrarily

$$F = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$$

$$u_1(0) = -\frac{u_2(0)}{2} - \frac{3}{14}x(0)$$

$$u_1(1) = u_2(1) + \frac{x(0)}{14}.$$

Substitute in J_2 and carry out the minimization. We obtain

$$u_2(0) = \frac{3}{7}x(0)$$

$$u_2(1) = -\frac{x(0)}{7}$$

as expected. In the second case, even though u_1 does not enter J_2 directly, it does affect J_2 through x.

- Information structure (b)

Now we examine RSP when only the values of the state variables are available to the leader. We shall see that the solvability becomes very stringent. (For generic solvability weneed $N \geq \frac{2n}{m_1}+1$, $N=2$, $\frac{2n}{m_1}+1=3$, thus the example here is in fact generically unsolvable.)

Consider now the representation

$$u_1 = K(x-(D+H_1G_1^o+H_2G_2^o)x(0))+G_1x$$

where K solves

$$(K+G_1)(H_2+H_1F) = F.$$

F is the same as in the last section. Let

$$K = \begin{bmatrix} k_1 & 0 & 0 \\ k_2 & k_3 & 0 \end{bmatrix}$$

$$k_1 = \text{arbitrary} \quad f_1=f_3=0$$

$$k_2 = \text{arbitrary}$$

$$k_3 = \frac{1}{2}-f_2.$$

k_1,k_2 do not enter into the solution, since u_1 cannot deduce any information of u_2 from x(0), the dependence of x(0) has not consequence to the solution. The same reasoning tells us that a penalty on $u_1(0)$ will also have no effect on the solution.

Note that given (a,b,c,d) f_3 is determined uniquely. Therefore, $f_3=0$ is a strict requirement on the problem (if $d\neq\frac{1}{2}$, the problem has no solution). This coincides with the statement before that since $N \not\geq \frac{2n}{m_1}+1$, the problem is generically unsolvable. Here we proceed with the assumption $d = \frac{1}{2}$ in order to verify that the strategy does indeed enforce the team solution.

- Let c=1, b=0, then $f_2 = -2$

$$K = \begin{bmatrix} k_1 & 0 & 0 \\ k_2 & \frac{5}{2} & 0 \end{bmatrix}$$

$$u_1(1) = \frac{11}{14} x(0) - 2u_2(0)$$

$$u_2(0) = \frac{3}{7} x(0)$$

$$u_2(1) = -\frac{1}{7} x(0).$$

- Let $c = \frac{1}{2}$, b=0, then $f_2 = 1$

$$K = \begin{bmatrix} k_1 & 0 \\ k_2 & -\frac{1}{2} \end{bmatrix}$$

$$u_1(1) = -\frac{x(0)}{2} + u_2(0)$$

$$u_2(0) = \frac{3}{7} x(0)$$

$$u_2(1) = -\frac{1}{7} x(0).$$

3.3. Behavior of Leader's Cost Under Large Threat

A natural question to pose after obtaining the results of the previous section is what can be done when there exists no solution to the set of conditions stated. It will be seen in this section that under certain mild conditions, infinite threat from the leader (i.e., leader threatens to drive the follower's cost to infinity if the follower does not perform as desired) can achieve the team solution for the leader. It therefore seems promising that perhaps near-team cost can be attained by using a very large but finite threat when the previously stated conditions are not satisfied. However, this expectation will be shown to be false in general if the leader's representation is continuous in u_2. No matter how large (but finite) the leader wants to penalize the follower's deviation, he cannot achieve arbitrary closeness to his team cost.

3.3.1. Solvability of RSP Under Infinite Threat

We shall be concerned with linear representation of the leader's strategy only. We study the solvability of RSP when the threat in the leader's strategy is weighted by a gain that tends to infinity. It is shown that under some mild conditions RSP is solved.

Without loss of generality (in the class of deterministic, centralized information structures), we assume information structure (a). Assume we adopt the representation (3.30) for the leader's strategy

$$u_1(u_2) = F(u_2 - G_2^o x_o) + G_1^o x_o,$$

and assume the optimal strategy of the follower, given that the leader has announced his strategy, is

$$u_2^* = G_2^o x_o + \Delta G x_o, \tag{3.33}$$

where $(G_1^o x_o, G_2^o x_o)$ is the team solution pair. From (3.18),

$$p[F'(H_1'Q_2 x + R_{21}u_1) + H_2'Q_2 x + R_{22}u_2] = 0 \tag{3.34}$$

$$u_1(u_2^*) = F\Delta G x_o + G_1^o x_o \tag{3.35}$$

$$x = Dx_o + H_1u_1(u_2^*) + H_2u_2$$

$$= Dx_o + H_1G_1^ox_o + H_2G_2^ox_o + H_1F\Delta Gx_o + H_2\Delta Gx_o$$

$$= (D + H_1G_1^o + H_2G_2^o)x_o + (H_1F + H_2)\Delta Gx_o \tag{3.36}$$

$$= x^t + \Delta x. \tag{3.37}$$

Note that

$$G_i^ox_o = G_ix^t \quad \text{(by the definition of } G_i\text{)}. \tag{3.38}$$

Rewrite (3.33), (3.35) using (3.38) and substitute together with (3.37) into (3.34)

$$p[(F'(H_1'Q_2+R_{21}G_1) + H_2'Q_2+R_{22}G_2)x^t + F'(H_1'Q_2\Delta x+R_{21}F\Delta Gx_o)$$
$$+ H_2'Q_2\Delta x + R_{22}\Delta Gx_o] = 0$$

or

$$\{[F'(H_1'Q_2+R_{21}G_1) + H_2'Q_2+R_{22}G_2)(D + H_1G_1^o+H_2G_2^o)] + [F'(H_1'Q_2(H_1F+H_2) + R_{21}F)$$
$$+ H_2'Q_2(H_1F+H_2) + R_{22}]\Delta G\}x_o = 0 \qquad \forall x_o.$$

Since x_o can be any vector in R^n, we have

$$\{[(F'(H_1'Q_2+R_{21}G_1)+H_2'Q_2+R_{22}G_2)(D+H_1G_1^o+H_2G_2^o)]+[R_{22}+H_2'Q_2H_2+F'H_1Q_2H_1F$$
$$+ F'R_{21}F+F'H_1'Q_2H_2+H_2'Q_]H_1F]\Delta G\} = 0 \tag{3.39}$$

$$G = -[(R_{22}+H_2'Q_2H_2)+F'(H_1'Q_2H_1+R_{21})F + F'H_1'Q_2H_2+H_2'Q_2H_1F]^{-1}[(F'(H_1'Q_2$$
$$+R_{21}G_1) + H_2'Q_2+R_{22}G_2)(D+H_1G_1^o+H_2G_2^o)]. \tag{3.40}$$

If F satisfies (3.27), then $\Delta G = 0$, and the leader's team solution is enforced. If there exists no F satisfying (3.27), the team solution is still attainable by the following.

Proposition 3.5

If $F \cdot 0 = 0 \quad \forall F$

$$F \to \infty$$

and

$$(H_1'Q_2H_1+R_{21})$$

is nonsingular, then the representation of the leader's strategy as in (3.30) will force the follower to adopt the corresponding team strategy.

Proof: Let $F \to \infty$, then along any direction in the $R^{m_1 \times m_2}$ space the denominator is of $0(\|F\|^2)$ and the numerator is of $0(\|F\|)$. Therefore, $\Delta G \to 0$ componentwise, $\Delta G = 0$ implies the leader's team solution is enforced. Q.E.D.

The above result is theoretically useful since it says that RSP is always solvable for this information structure provided that infinite gain is possible. However, the infinite threat is not physically realizable, therefore, it is natural to ask whether the team cost can be approached arbitrarily close given a finite gain that is large enough.

3.3.2. Effect of Finiteness of Threat

It is shown in this section that if we consider F not identically equal to infinity, the largeness of F will not enable the leader to approach team cost arbitrarily. A key assumption in Proposition 3.5 is that $F \cdot 0 = 0$, which means that even though F is an infinite threat, if the follower plays team exactly, the threat will have no effect. However, for the case F being finite (no matter how large), the follower's decision cannot be made exactly team (first order condition in Section 3.2 is assumed not satisfied). The deviation can be shown $\sim 0(\|F\|^{-1})$, which is then amplified by F. Therefore, there will be a sizable deviation in the leader's cost.

We first examine the effect of control offsets to the leader's cost

$$u_1 = G_1^o x_o + F\Delta G x_o$$
$$u_2 = G_2^o x_o + \Delta G x_o. \tag{3.40}$$

Then

$$x = \Delta x_o + H_1 G_1^o x_o + H_2^o G_2^o x_o + H_1 F\Delta G x_o + H_2 \Delta G x_o$$
$$= x^t + (H_1 F + H_2)\Delta G x_o \tag{3.41}$$

$$J_1(u_1,u_2) = (x^t+\Delta x)'Q_1(x^t+\Delta x) + (u_1^t+\Delta u_1)'R_{11}(u_1^t+\Delta u_1)$$
$$+ (u_2^t+\Delta u_2)'R_{12}(u_2^t+\Delta u_2)$$
$$= J_1^t + \Delta J_1 \tag{3.42}$$

$$\Delta J_1 = x_o'(\Delta G'[(H_1F+H_2)'Q_1(D+H_1G_1^o+H_2G_2^o) + F'R_{11}G_1^o + R_{12}G_2^o]$$
$$+ [(D+H_1G_1^o+H_2G_2^o)'Q_1(H_1F+H_2) + G_1^{o\prime}R_{11}F + G_2^{o\prime}R_{12}]\Delta G$$
$$+ \Delta G'[(H_1F+H_2)'Q_1(H_1F+H_2) + F'R_{11}F + R_{12}]\Delta G)x_o. \tag{3.43}$$

If the leader's cost is continuous with respect to $\|F\|$, letting $F\to\infty$ in (3.43) should imply $\Delta J_1 \to 0$. However, we will see that in general it is not true by deriving $\lim_{\|F\|\to\infty} \Delta J_1$. $A_o F\to\infty$, we retain the dominant terms only

$$\therefore \quad \Delta G \sim -(F'(H_1'Q_2H_1 + R_{21})F)^{-1}F'(H_1'Q_2 + R_{21}G_1) \tag{3.44}$$

$$\Delta J_1 \sim \Delta G'F'((H_1'Q_1H_1+R_{11})G_1^o + H_1'Q_1(D+H_2G_2^o)$$
$$- (H_1'Q_1H_1+R_{11})F(F'(H_1'Q_2H_1+R_{21})F)^{-1}F'(H_1'Q_2+R_{21}G_1)) \tag{3.45}$$

where $\ker F = 0$ has been assumed (generically true if $m_1 \geq m_2$).

Since ker $F = 0$ orthogonal matrices U,V

$$F = U \begin{bmatrix} \tilde{F} \\ 0 \end{bmatrix} V \tag{3.46}$$

$$\therefore \; U'FV' = \begin{bmatrix} \tilde{F} \\ 0 \end{bmatrix}$$

where $\tilde{F}$ is nonsingular. Then, let $R = H_1'Q_1H_1 + R_{21} > 0$,

$$F(F'RF)^{-1}F' = F(V'[\tilde{F}' \;\; 0]U'RU \begin{bmatrix} \tilde{F} \\ 0 \end{bmatrix} V)^{-1}F'.$$

Let

$$U'RU = \begin{bmatrix} R_1 & R_2 \\ R_2' & R_3 \end{bmatrix} \quad (\|U'RU\| = \|R\|) \tag{3.47}$$

$$\begin{aligned} F(F'RF)^{-1}F &= FV'(\tilde{F}'R_1\tilde{F})^{-1}VF' \\ &= U(U'FV')\tilde{F}^{-1}R_1^{-1}\tilde{F}'^{-1}(VF'U)U' \\ &= U \begin{bmatrix} \tilde{F} \\ 0 \end{bmatrix} \tilde{F}^{-1}R_1^{-1}\tilde{F}'^{-1}[\tilde{F}' \;\; 0]U' \\ &= UR_1^{-1}U' \end{aligned} \tag{3.48}$$

$$\Delta G'F' = -H_1'Q_2 + R_{21}G_1)'F(F'(H_1'Q_2H_1 + R_{21})F)^{-1}F'.$$

As $F \to \infty$

$$\Delta G'F' \to -(H_1'Q_2 + R_{21}G_1)'UR_1^{-1}U' \quad \text{as derived before} \tag{3.49}$$

$$\begin{aligned} \Delta J_1 \to -(H_1'Q_2 + R_{21}G_1)'UR_1^{-1}U'[&(H_1'Q_1H_1+R_{11})G_1^o + H_1'Q_1(D+H_2G_2^o) \\ &- (H_1'Q_1H_1 + R_{11})UR_1^{-1}U'(H_1'Q_2 + R_{21}G_1)] \end{aligned} \tag{3.50}$$

which in general is nonzero, thus proving the asserted result.

It is certainly of importance to investigate in the case of finite threat and failure of the first order condition whether there exists a near optimal strategy for the leader. One method is to assume one representation (selected from the linear class) and perform parameter optimization. A possible conjecture is that the minimum solution F in (3.27) will correspond to the nearest optimal solution. However, a verification of this conjecture is not yet available.

Note that from (3.50), ΔJ_1 will in fact tend to zero on a variety of the parameter space. Since we are only interested in this result when conditions like (3.27) fial, in some cases it may happen that this variety will have high probability of occurrence on the subset of the parameter space where (3.27) fails. However, it appears "generically" that ΔJ_1 tends to a nonzero limit for Stackelberg strategy with very large threat.

It should be noted also that the conclusion drawn here is for u_1 as a continuous function of u_2. If u_1 is allowed to be discontinuous, ΔJ_1 will in fact be zero for finite threats that are large enough.

3.3.3. Examples

We use the example in Section 3.2.3 to illustrate the effect of infinite and finite threats. It is shown that if each component of the threat tends to infinity at equal rate when the leader announces his strategy, then the team solution is indeed enforced. However, if the threat coefficients tend to infinity (at equal rate) in the leader's cost, the limiting cost is shown to be higher than the team cost.

From Secion 3.2.3, we have the following representation for $u_1(0)$ and $u_1(1)$

$$u_1(0) = f_1(u_2(0) - \frac{3}{7}x_o) - \frac{3}{7}x_o$$

$$u_1(1) = f_2(u_2(0) - \frac{3}{7}x_o) + f_3(u_2(1) + \frac{1}{7}x_o) - \frac{1}{14}x_o$$

where f_1, f_2, f_3 are coefficients in the threat matrix. Then,

$$x(1) = \frac{4}{7}x_o + f_1(u_2(0) - \frac{3}{7}x_o) - u_2(0)$$

$$x(2) = \frac{15}{14}x_o + 2f_1(u_2(0) - \frac{3}{7}x_o) + f_2(u_2(0) - \frac{3}{7}x_o) + f_3(u_2(1) + \frac{1}{7}x_o) - 2u_2(0) + u_2(1).$$

Let

$$u_2(0) = \frac{3}{7}x_o + \Delta g_o x_o$$

$$u_2(1) = -\frac{1}{7}x_o + \Delta g_1 x_o.$$

Then

$$u_1(0) = (f_1\Delta g_o - \frac{3}{7})x_o$$

$$u_1(1) = (f_2\Delta g_o + f_3\Delta g_1 - \frac{1}{14})x_o.$$

When f_1, f_2, f_3 all tend to $+\infty$ at equal rate, asymptotically

$$\Delta g_o \sim \frac{2(1+b)f_2f_3-(1+3b-3a-3ab)f_1f_3}{14[(1+b)f_3((5+a)f_1^2+(1+b)f_2^2+4f_1f_2)-f_3(2f_1+(1+b)f_2)^2]} \to 0$$

$$\Delta g_1 \sim \frac{-2((1+b)f_2+2f_1)(1+b)f_2-2(2-3a)f_1)+(b-1)((5+a)f_1^2+(1+b)f_2^2+4f_1f_2}{14[(1+b)f_3((5+a)f_1^2+(1+b)f_2^2+4f_1f_2)-f_3(2f_1+(1+b)f_2)^2]} \to 0.$$

Thus, when $f_1 = f_2 = f_3 = +\infty$

$$\Delta g_o = \Delta g_1 = 0$$

and the team solution is enforced by the leader.

If we retain f_1, f_2, f_3 and substitute the expressions into J_1, we get

$$J_1 = J_1^t + \Delta J_1$$

where J_1^t is the team cost and ΔJ_1 is the deviation due to Δg_o and Δg_1. As $f_1, f_2, f_3 \to \infty$ at equal rate,

$$\Delta J_1 \sim (10f_1^2+8f_1f_2+4f_2^2)\Delta g_o^2 + 4f_3^2\Delta g_1^2 + (8f_2f_3+8f_1f_3)\Delta g_o \Delta g_1 .$$

From the above, we know $\Delta g_o, \Delta g_1 \sim 0(\frac{1}{f_i})$. But the quadratic-$f_i$ coefficients make each term tending to a finite limit. If

$$\lim_{\substack{f_i \to \infty \\ f_j}} f_i/f_j = 1 \qquad \forall i,j \in \{1,2,3\}$$

$$x(1) = (\frac{1}{7} + (f_1-1)\Delta g_o)x_o$$

$$x(2) = (\frac{1}{14} - 2\Delta g_o + \Delta g_1 + 2f_1\Delta g_o + f_2\Delta g_o + f_3\Delta g_1)x_o .$$

Substitution of these expressions into J_2 and minimization with respect to $u_2(0)$ and $u_2(1)$ render

$$\frac{(6c-4)+(4-6a)f_1-(1+b)f_2}{14} + ((5+a)f_1^2 + ((1+b)f_2^2 + 4f_1f_2 - 10f_1 - 4f_2 + 5 + c)\Delta g_o$$
$$+ (2f_1f_3 + (1+b)f_2f_3 - 2f_3 + 2f_1 + f_2 - 2)\Delta g_1 = 0$$

and

$$\frac{(1-b)f_3+(1-2d)}{14} + ((1+b)f_2f_3 + 2f_1f_3 + 2f_1 + f_2 - 2f_3 - 2)\Delta g_o$$
$$+ ((1+b)f_3^2 + 2f_3 + (1+d))\Delta g_1 = 0.$$

When f_1, f_2, f_3 are chosen to satisfy the sufficiency conditions derived in Section 3.3.2, namely,

$$(6c-4) + (4-6a)f_1 - (1+b)f_2 = 0$$

$$(1-b)f_3 + (1-2d) = 0,$$

then $$\Delta g_o = \Delta g_1 = 0$$

and the team solution is enforced

$$\lim_{f_i \to \infty} \Delta J_1 = \frac{1}{(1+a+5b+ab)^2} [22(4b+6a+6ab)^2 + 4(8-b^2-37a-3b+ab)^2$$
$$+ 15(8-b^2-37a-3b+ab)(4b+ba+6ab)].$$

Thus, we conclude that the team cost cannot be approached arbitrarily close with large threats.

It has been mentioned in the previous section that we should examine the continuity of J_1 only for the cases the sufficiency conditions in Section 3.2 fail. In this example, they occur at $a = \frac{2}{3}$ or $b = 1$. For $b = 1$

$$\Delta J_1 \to \frac{1}{(6+2a)^2} [352(1+3a)^2 + 64(1-9a)^2 + 256(1-9a)(1+3a)].$$

3.4 Decentralized and Stochastic RSP

3.4.1. Introduction

The cases we shall examine here are the deterministic decentralized and stochastic state feedback information structures. Due to the difficulty of the general decentralized team problem (as will be explained later), the team problem is solved under the restriction of linear strategies. Sufficient conditions similar to those obtained in section 3.2 can then be stated, but, as to be expected, they become slightly more stringent. In the stochastic case, the problem is not solved in general; it is only after some further restrictions are imposed on the information structure that non-void conditions for RSP can be obtained.

Intuitively, RSP can be solved in two ways. One is to use the infinite threat concept discussed in section 3.3. The other is to alter the follower's objective function so that the optimal follower strategy coincides with the team strategy. The former method meets with difficulties in both decentralized and noisy state information cases. In the first case, the leader can only enforce the team trajectory projected onto his observation space, which in general does not imply that his team cost is attained. In the second case, the leader is unable to implement the threat term (that vanishes upon the enforcement of the team solution) due to the random nature of the state trajectory. The latter method, however, can be applied to the deterministic decentralized RSP provided linear representations of the strategies are constrained. But the method still fails in the stochastic case. Therefore, we allow the leader to have access to the follower's past control, the problem then reduces to the same framework as before.

3.4.2. Decentralized RSP

We approach RSP under information structure (d) in the same manner as in section 3.2. However, now the team solution is not as easily obtainable as in the previous case. It is known that the general team solution under the decentralized information pattern is in general non-linear, and, in fact, the problem is not always analytically solvable ([15], [16]). The difficulty lies in the dual role of the control variables, namely, control and estimation. Specifically, if the projection approach used in section 3.2 is used here, the projections of the state variables onto the observation space cannot be evaluated since the distribution of the states are affected by the past controls which in turn depend on the projection of the states. Therefore, here we constrain the strategies to be linear in observation. The optimization of leader's performance is carried out under this constraint by using the parameter optimization technique (the discrete-time and finite-horizon equivalent of the continuous time approach in [17]). The leader then tries to enforce this solution. Once linearity is assumed, sufficient conditions for RSP solutions can be derived in the same way as in section 3.2, but, as expected, these conditions are more stringent than the centralized, deterministic counterparts.

Team Solution Under Linear Representation Constraint

In this section, we use the parameter optimization technique to obtain the best linear solution for the leader's team problem. The decentralized outputs are assumed to be linear functions of the past states, i.e.,

$$Y_i\ (k) = \sum_{j=0}^{k} C_{ij}\ (k)\ X\ (j) \qquad \in \mathbb{R}^{Yi} \tag{3.51}$$

is the observation vector for player i at stage k. The states may be noise-corrupted.

Convert the observations to static form in the usual way.

$$Y_i = C_i X \tag{3.52}$$

where C_i is a block lower triangular matrix.

The objective is to solve

$$\min J_1 = E\ \{X'Q_1\ X + U_1'R_{11}U_1 + U_2'R_{12}U_2\} \tag{3.53}$$

such that

$$U_i = G_i Y_i \qquad i = 1,2$$

G_i is a block lower triangular matrix.

The following proposition states the sufficient conditions for the above problem.

<u>Proposition 3.6</u>

If there exists a unique quadruplet (G_1, G_2, P, Λ) satisfying

$$(I-H_1\ G_1\ C_1 - H_2G_2C_2)\ P\ (I-H_1G_1C_1 - H_2G_2C_2)' = D\ \Sigma_o\ D' \tag{3.54}$$

$$(Q_1 + C_1'G_1'R_{11}G_1C_1 + C_2'G_2'R_{12}G_2C_2)$$

$$+ (I-H_1G_1C_1 - H_2G_2C_2)'\ \Lambda\ (I-H_1G_1C_1 - H_2G_2C_2) = 0 \tag{3.55}$$

$$ut_1\ \{\ C_1P\ (C_1'G_1'R_{11} - (I-H_1G_1C_1 - H_2G_2C_2)'\Lambda\ H_1)\ \} = 0 \tag{3.56}$$

$$ut_2 \{ C_2 P (C_2' G_2' R_{12} - (I - H_1 G_1 C_1 - H_2 G_2 C_2)' \Lambda H_2) \} = 0 \qquad (3.57)$$

(ut_i $(\cdot)$ $\triangleq$ block upper triangular portion of the matrix with block dimension $\Upsilon_i \times M_i$) (3.58)

$$R_{12} + H_2' \Lambda H_2 > 0 \qquad (3.59)$$

$$R_{11} + H_1' \Lambda H_1 > 0 \qquad (3.60)$$

$$P > 0$$

(where $\Sigma_o = E\,[X_o X_o']$)

then

$$U_1 = G_1 Y_1 \qquad (3.61)$$

$$U_2 = G_2 Y_2 \qquad (3.62)$$

solve the problem (3.53)

Proof: See Appendix II. Q.E.D.

Conditions for Enforcing the Team Solution

We now derive the sufficient conditions for the leader to enforce his best linear decentralized team solution. The development is similar to the centralized case, in fact, some of the previous results are directly applicable here.

Theorem 3.7

Let $Z_i = \{ Y_i(0), \text{--}, Y_i(k) \}$

Assume

(i) $\gamma(Z_1)$ is a causal, differentiable, class-T function

(ii) $J_2(\gamma(Z_1), u_2)$ is convex over the convex set $U_2 = \{u_2 | u_2(k)$ measurable with respect to $\{y_2(0), \ldots, y_2(k)\}\}$

(iii) $\exists F \ni$

F solves

$$[(R_{22} G_2 C_2 + H_2' Q_2) + F'(R_{21} G_1 C_1 + H_1' Q_2)] (I - H_1 G_1 C_1 - H_2 G_2 C_2)^{-1} D = 0$$

and

$$F = \nabla_{y_1} \gamma(y_1) | C_1 (H_1 F + H_2)$$

then $\quad \gamma_1 = \gamma_1^t$

$y_1 = \gamma(Z_1)$ will force u_2 to play u_2^t

<u>proof</u>:

Assumptions (i), (ii) guarantee that the minimization of $J_2(\gamma(Z_1), u_2)$ results in $u_2 = u_2^t$, then (u_1^t, u_2^t) is a global minimum of $J_2(\gamma(Z_1), u_2)$. Therefore, it suffices to show that $P_2[\nabla_{u_2} J_2]| = 0$ implies condition (iii).

$$P_2[\nabla_{u_2} J_2] = P[(\nabla_{u_2} X)' Q_2 X + (\nabla_{u_2} \gamma)' R_{21} \gamma + R_{22} u_2] \tag{3.65}$$

$$\nabla_{u_2} X = H_2 + H_1 \nabla \gamma \tag{3.66}$$

$$\nabla_{u_2} \gamma = \nabla_{y_1} \gamma(y_1) \nabla_{u_2} y_1$$

$$= \nabla_{y_1} \gamma(y_1) C_1 (H_2 + H_1 \nabla_{u_2} \gamma) \tag{3.67}$$

(3.65) becomes

$$P[(\nabla_{u_2} \gamma(y_1))' (H_1' Q_2 X + R_{21} \gamma(y_1)) + H_2' Q_2 X + R_{22} u_2] = 0 \tag{3.68}$$

When $u_2 = G_2 y_2 = G_2 C_2 X$

$\gamma(y_1) = G_1 C_1 X$

(3.68) becomes

$$P \{ [(\nabla_{u_2} \gamma(y_1))' \mid (H_1'Q_2 + R_{21}G_1C_1) + H_2'Q_2 + R_{22}G_2C_2] \, X^t \} = 0$$

$$X^t = (I - H_1G_1C_1 - H_2G_2C_2)^{-1} D \, X_o$$

It is sufficient that

$$[(\nabla_{u_2} \gamma(y_1))' \mid_{y_1 = y_1^t} (H_1'Q_2 + R_{21}G_1C_1) + H_2'Q_2 + R_{22}G_2C_2] (I - H_1G_1C_1 - H_2C_2)^{-1} D = 0 \qquad (3.69)$$

Let $\nabla_{u_2} \gamma(y_1) \mid_{y_1 = y_1^t} = F$

Substitute in (3.67) and (3.69), result follows.

It is seen that the result is almost identical to that of the full state information case Therefore, all the qualitative discussion pertaining to that case carries over here. The specialization to linear strategy is straight forward, therefore is omitted here.

3.43 Stochastic RSP:

As mentioned in the introduction, stochastic RSP with state information only is not solvable due to the randomness of the team trajectory. The only information structure that can be shown solvable under the stochastic setting is the one including the perfect knowledge of follower's action. This assumption exactly bypasses the difficulty since u_2 is known and can be used to check against u_2^t. Once this assumption is made, the derivation becomes almost identical as in section 3.2.

For LQ setting given state information, the separation theorem holds, therefore, the team optimal solutions are as in (3.9)

$u_i = G_i \, X$

where G_i is block diagonal.

We can then immediately state the sufficient conditions.

Theorem 3.8

Given information structure (e)

Assume

(i) (Z_1) is a causal, differentiable, class-T function

(ii) $J_2\ (\gamma(Z_1),\ u_2)$ is convex over the convex set $U_2 = \{u_2 |\ u_2\ (k)$ measurable with respect to $\{X(0), \ldots, X(k)\}$

(iii) $\exists\, F \ni$

$$[F'(H_1'Q_2 + R_{21}G_1) + (H_2'Q_2 + R_{22}G_2)]\ (I - H_1G_1 - H_2G_2)^{-1}\ [D \quad I] = 0 \tag{3.70}$$

and

$$\nabla_X^{(1)}\ \gamma(X,\ u_2)\big|\ (H_2 + H_1\ F) + \nabla_{u_2}^{(2)}\ \gamma(X,\ u_2)\big|\ = F \tag{3.71}$$

Then

$u_1 = \gamma\ (Z_1)$ will force u_2 to adopt u_2^t

Proof:

Using the proof of Theorem 3.2, we get

$$[(\nabla_{u_2}\gamma(X,\ u_2)\big|_{\substack{u_2 = u_2^t \\ X = X^t}})'(H_1'\ Q_2 + R_{21}G_1) + (H_2'Q_2 + R_{22}G_2)]\ X^t = 0$$

$$X^t = D\ X_o + H_1G_1\ X^t + H_2G_2\ X^t + W$$

$$X^t = (I - H_1G_1 - H_2G_2)^{-1}\ [D \quad I] \begin{bmatrix} X_o \\ W \end{bmatrix}$$

It is sufficient that

$$[(\nabla_{u_2}\ \gamma(X,\ u_2)\big|_{\substack{u_2 = u_2^t \\ X = X^t}})'(H_1'Q_2 + R_{21}G_1) + (H_2'Q_2 + R_{22}G_2)](I - H_1G_1 - H_2G_2)^{-1}[D \quad I] = 0 \tag{3.72}$$

Now

$\nabla_{u_2} \gamma(X, u_2)$ satisfies (from (3.21))

$$\nabla_X^{(1)} \gamma(X, u_2)\Big|_{\substack{X = X^t \\ u_2 = u_2^t}} \left(H_2 + H_1 \nabla_{u_2} \gamma(X, u_2)\Big|_{\substack{X = X^t \\ u_2 = u_2^t}}\right) \nabla_{u_2}^{(2)} \gamma(X, u_2)\Big|_{\substack{X = X^t u_2 \\ u_2 = u_2}} = \nabla_{u_2} \gamma(X, u_2)\Big|_{\substack{X = X^t \\ u_2 = u_2^t}}$$

Let $F = \nabla_{u_2} \gamma(X, u_2)\Big|_{\substack{X = X^t \\ u_2 = u_2^t}}$, we have the stated result

In (3.70), we have $\frac{N(N + 1)}{2} m_1 \times m_2$ unknowns in F (F is causal), and N $(2n \times m_2)$ equations. Then if $N \geq \frac{4n}{m_1} - 1$, F is generically solvable.

The stochastic RSP is still an open problem. Even though we know that under the state information solution, solution does not exist. It will be of a great deal of interest to see how near-optimal is the parameter optimization approach. The near-optimality of some intuitive method, such as the use of best team state trajectory estimate or just the plain certainty equivalence, should also be investigated.

4. CONCLUSION

In this report, we have studied the application of dynamic-to-static conversion technique to the Restricted Stackelberg Problem. RSP is a restricted version of the Stackelberg equilibrium solution concept. It is an important modeling tool for the economic systems and large scale engineering systems.

The definitions of the team, Stackelberg, and restricted Stackelberg problems are first stated. We then give the precise statement of the problem under consideration and introduce the conversion technique which is the backbone of this analysis. The past work and results are briefly summarized and the contributions of this report and pointed out to close off Chapter 2.

The main results and discussions are presented in Chapter 3. RSP under five different information structures is considered. Three of the information structures are centralized, deterministic, the others are deterministic decentralized and stochastic. The deterministic centralized information patterns illustrate how RSP is approached and permits the examination of various qualitative aspects of its solution. They also show how the restriction on the information structure affects the solvability of RSP. The decentralized information pattern encounters a particular difficulty with regard to RSP, viz., in the solution of the corresponding team problem. Since the team solution is difficult to obtain in general, we settle for a suboptimal result, the best linear team solution. The sufficient conditions for the leader to enforce this solution are then derived. The stochastic information patterns create another difficulty in RSP, the inability to formulate the threat in the presence of random noise. This problem is bypassed by assuming the follower's past controls to be available to the leader.

The centralized cases are studied in detail. Sufficient conditions for RSP solutions are derived and how they are affected by the information structure of the leader is discussed. We then examine the stationarity and the convexity conditions for the general nonlinear representation of the leader's strategy to ensure that the team solution is indeed also the follower's optimal operating point. The result restricted to linear representations of the leader's strategy is then presented, motivated by the observation that nonlinearity does not add any significant advantage and poses difficulty in the convexity condition. An example is also presented to verify the derived results.

Noting that the sufficient conditions are not always satisfied, a natural query arises: in the case the stated conditions fail to be satisfied, can the leader attain a cost arbitrarily close to the team cost by choosing a threat as large as he desires (but finite)? To address this question it is found that if the threat is infinite, RSP is solved (under some mild conditions). However, if the threat is large but finite (no matter how large), in general there is always an offset, bounded away from zero, in his cost from the team cost. It should be noted that this assertion is not true if discontinuous strategies are allowed. This result, though reduces the hope of a continuous, guaranteed near optimal solution, does offer a design alternative if the offset is not very large. An example is also presented to verify the above result.

Strictly speaking, the general decentralized RSP is not solved. The problem lies in the fact that the corresponding team solution is not solved in general. However, if the structure of strategies is restricted to linear, then, by using parameter optimization, the team problem can be solved. RSP is solved the same way as the centralized case once the linearity assumption is adopted. The stochastic case with state information is not solved (and seems unsolvable in its full generality) due to the lack of redundant information to implement the threat (the state trajectory corresponds to a sample path of a random process). To bypass the problem, we allow the leader to have access to the follower's past controls. The problem then reduces to the deterministic case. The stochastic decentralized RSP with the leader having the follower's past controls, though not presented, can be tackled in the same manner as the combination of the above two problems. However, the linear representation constraint again has to be used. Suboptimal results may be obtained via parameter optimization, but are not pursued in this report.

The static conversion has proved invaluable in simplifying the conditions and the analysis of RSP. There certainly remain a great deal of open questions, even for this special type of problem. The suboptimal strategies need to be investigated in the deterministic case when the derived conditions are not satisfied, and in the stochastic case when the information is restricted to the past states only. The hierarchical result also needs to be developed (it will be an easy extension of the results stated here) because of the unique feature of RSP that the follower is under no protection from the leader's manipulation. The conditions for RSP solutions should be interpreted from a qualitative, perhaps geometric, point of view. Specific applications should also be investigated to demonstrate that RSP is not merely a theoretical pastime but has definite practical value.

REFERENCES

1. R. Radner, "Team Decision Problem," Ann. Math. Stat., Vol. 33, No. 3, pp. 851-881, 1962.

2. B. Bryson and Y. C. Ho, Applied Optimal Control, Wiley, New York, pp. 277-295, 1975.

3. A. W. Starr and Y. C. Ho, "Nonzero-Sum Differential Games," J. of Optimization Theory and Applications, Vol. 3, pp. 184-206, 1969.

4. C. I. Chen and J. B. Cruz, Jr., "Stackelberg Solution for Two-Person Games with Biased Information Patterns," IEEE Trans. on Automatic Control, Vol. AC-17, pp. 791-798, December 1972.

5. T. Basar and H. Selbuz, "Closed-Loop Stackelberg Strategies with Application in the Optimal Control of Multilevel Systems," IEEE Trans. on Automatic Control, Vol. AC-24, pp.

6. G. P. Papavassilopoulos and J. B. Cruz, Jr., "Sufficient Conditions for Stackelberg and Nash Strategies with Memory," J. of Optimization Theory and Applications, Vol. 31, No. 2, June 1980.

7. G. P. Papavassilopoulos, "Leader-Follower and Nash Strategies with State Information," Ph.D. thesis, University of Illinois, July 1979.

8. T. Basar, "Equilibrium Solutions in Two-Person Quadratic Decision Problems with Static Information Structures," IEEE Trans. on Automatic Control, Vol. AC-20, pp. 320-328, June 1975.

9. T. Basar, "Decentralized Multicriteria Optimization of Linear Stochastic Systems," IEEE Trans. on Automatic Control, Vol. AC-23, pp. 233-241, April 1978.

10. G. P. Papavassilopoulos, "Solution of Some Stochastic Quadratic Nash and Leader-Follower Games," submitted for publication.

11. Y. C. Ho and K. C. Chu, "Team Decision Theory and Information Structures in Optimal Control Problems - Part I," IEEE Trans. on Automatic Control, Vol. AC-17, pp. 15-22, February 1972.

12. H. von Stackelberg, The Theory of the Market Economy, Oxford University Press, Oxford, England, 1952.

13. Y. C. Ho, P. B. Luh, and R. Muralidharan, "Information Structure, Stackelberg Games, and Incentive Controllability," submitted for publication.

14. T. Basar, "Hierarchical Decision Making Under Uncertainty," in Dynamic Optimization and Mathematical Economics, Pan-Tai Liu, Ed., Planum Press, New York, pp. 205-221, 1980.

15. T. Basar, "On the Optimality of Nonlinear Designs in the Control of Linear Decentralized Systems," IEEE Trans. on Automatic Control, Vol. AC-21, p. 797, October 1976.

16. H. Witsenhausen, "A Counterexample in Stochastic Optimum Control," SIAM J. Control, Vol. 6, No. 1, pp. 131-147, 1968.

17. W. S. Levine and M. Athans, "On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems," IEEE Trans. on Automatic Control, Vol. AC-15, No. 1, pp. 44-48, February 1970.

18. I. B. Rhodes and D. G. Luenberger, "Differential Games with Imperfect State Information," IEEE Trans. on Automatic Control, Vol. AC-14, No. 1, pp. 29-38, February 1969.

19. I. B. Rhodes and D. G. Luenberger, "Stochastic Differential Games with Constrained State Estimators," IEEE Trans. on Automatic Control, Vol. AC-14, N .. 5, pp. 476-481, October 1969.

20. G. P. Papavassilopoulos and J. B. Cruz, Jr., "Nonclassical Control Problems and Stackelberg Games," IEEE Trans. on Automatic Control, Vol. AC-24, No. 2, pp. 155-166, April 1979.

21. B. Tolwinski, "Closed-Loop Stackelberg Solution to Multi-Stage Linear-Quadratic Games," to appear in J. of Optimization Theory and Applications.

APPENDIX I

PROOF OF PROPOSITION 3.1

We prove it by induction

1. At stage N-1, calculate $u_i(N-1)$ in terms of $x(N-1)$

$$u_i(N-1) = -R_{ii}^{-1}(N-1)B_i'(N-1)Q_i(N)x(N)$$

$$= -R_{ii}^{-1}(N-1)B_i'(N-1)Q_i(N)(A(N-1)x(N-1) + B_1(N-1)u_1(N-1) + B_2(N-1)u_2(N-1))$$

$$\therefore \begin{bmatrix} I-d_{N-1,N}^{(1)}B_1(N-1) & -d_{N-1,N}^{(1)}B_2(N-1) \\ -d_{N-1,N}^{(2)}B_1(N-1) & I-d_{N-1,N}^{(2)}B_2(N-1) \end{bmatrix} \begin{bmatrix} u_1(N-1) \\ u_2(N-1) \end{bmatrix} = \begin{bmatrix} d_{N-1,N}^{(1)} \\ d_{N-1,N}^{(2)} \end{bmatrix} A(N-1)x(N-1).$$

By assumption,

$$\begin{bmatrix} u_1(N-1) \\ u_2(N-1) \end{bmatrix} = \begin{bmatrix} g_1(N-1) \\ g_2(N-1) \end{bmatrix} A(N-1)x(N-1).$$

2. Assume similar procedure can be carried out to obtain

$$u_i(j) = g_i(j)x(j) \qquad \text{for } j = k+1,\dots,N-1.$$

Then,

$$x(\ell) = \prod_{i=k+1}^{\ell-1}(A(i) + B_1(i)g_1(i) + B_2(i)g_2(i))x(k+1))$$

$$u_i(k) = -\sum_{\ell=k+1}^{N} R_{ii}^{-1}(k)B_i'(k)\Phi'(\ell-1,k)Q_i(\ell)x(\ell)$$

$$= [-\sum_{\ell=k+1}^{N} R_{ii}^{-1}(k)B_i'(k)\Phi'(\ell-1,k)Q_i(\ell)\prod_{i=k+1}^{\ell-1}(A(i) + B_1(i)g_1(i) + B_2(i)g_2(i))]x(k+1)$$

$$= d_{k,k+1}^{(i)}x(k+1).$$

We can now apply the same reduction as stage N-1, and the result follows by induction. Q.E.D.

APPENDIX II

PROOF OF PROPOSITION 3.6

Substitute $u_i = G_i Y_i$ into (3.53)

$$J_1 = E\{x'(Q_1 + C_1'G_1'R_{11}G_1C_1 + C_2'G_2'R_{12}G_2C_2)x\}$$

$$= tr\{(Q_1 + C_1'G_1'R_{11}G_1C_1 + C_2'G_2'R_{12}G_2C_2)E[x\ x']\}.$$

Substitute $u_i = G_i Y_i$ into the state equation, then

$$(I-H_1G_1C_1-H_2G_2C_2)x = Dx_o.$$

Therefore,

$$(I-H_1G_1C_1-H_2G_2C_2)E[x\ x'](I-H_1G_1C_1-H_2G_2C_2)' = DE[x_o x_o']D'$$

where

$$DE[x_o x_o']D' = D\Sigma_o D'$$

is assumed known.

Let $P = E[x\ x']$ and use matrix Lagrange multiplier, we have converted the problem to one that chooses G_1, G_2, P, Λ to minimize

$$L(G_1,G_2,P,\Lambda) = tr[(Q_1 + C_1'G_1'R_{11}G_1C_1 + C_2'G_2'R_{12}G_2C_2)P$$
$$+ \Lambda((I-H_1G_1C_1-H_2G_2C_2)P(I-H_1G_1C_1-H_2G_2C_2)'-D\Sigma_o D')].$$

Set

$$\frac{dL}{d\varepsilon}(G_1,G_2,P,\Lambda+\varepsilon\Delta\Lambda)\Big|_{\varepsilon=0} = 0 \qquad \forall\Delta\Lambda \in \mathbb{R}^{(N+1)n\times(N+1)n}$$

we get

$$(I-H_1G_1C_1-H_2G_2C_2)P(I-H_1G_1C_1-H_2G_2C_2)' = D\Sigma_o D'. \qquad (3.54)$$

Set

$$\frac{dL}{d\varepsilon}(G_1,G_2,P+\varepsilon\Delta P,\Lambda)\Big|_{\varepsilon=0} = 0 \qquad \forall\Delta\rho \in \mathbb{R}^{(N+1)n\times(N+1)n}$$

we get

$$(Q_1 + C_1'G_1'R_{11}G_1C_1 + C_2'G_2'R_{12}G_2C_2) + (I-H_1G_1C_1-H_2G_2C_2)'\Lambda(I-H_1G_1C_1-H_2G_2C_2) = 0 \tag{3.55}$$

Set

$$\frac{dL}{d}(G_1 + \varepsilon\Delta G_1, G_2, P, \Lambda)\Big|_{\varepsilon=0} = 0 \qquad \forall \Delta G_1 \in \text{lower block triangular } R^{Nm_1 \times Nr_1}$$

matrices with each block of dimension $m_1 \times r_1$

$$\mathrm{tr}\{C_1P(C_1'G_1'R_{11} - (I-H_1G_1C_1-H_2G_2C_2)'\Lambda H_1\Delta G_1\} = 0.$$

Let

$$K = C_1P(C_1'G_1'R_{11} - (I-H_1G_1C_1-H_2G_2C_2)'\; H_1 = \begin{bmatrix} k_{11} & \cdots & k_{1N} \\ \vdots & & \vdots \\ k_N & \cdots & k_{NN} \end{bmatrix}$$

where $k_{ij} = r_1 \times m_1$ block

$$\Delta G_1 = \begin{bmatrix} \Delta G_{11} & & 0 \\ \vdots & \ddots & \\ \Delta G_{N1} & \cdots & \Delta G_{NN} \end{bmatrix}$$

$\Delta G_{ij} = m_1 \times r_1$ block

$$\mathrm{tr}K\Delta G = [k_{11} \;\cdots\; k_{1N}] \begin{bmatrix} \Delta G_{11} \\ \vdots \\ \Delta G_{N1} \end{bmatrix} + \cdots + [k_{N-1,N-1} k_{N-1,N}] \begin{bmatrix} G_{N-1,N-1} \\ G_{N,N-1} \end{bmatrix} + k_{NN}\Delta G_{NN}$$

$$= 0 \qquad \forall \Delta G_{ij} \in R^{m_1 \times r_1}$$

$$[k_{11} \quad k_{1N}] = 0$$

$$[k_{N-1,N-1} k_{N-1,N}] = 0$$

$$[k_{NN}] = 0.$$

Or,

$$ut_1[K] = 0.$$

Similarly for G_2, we have

$$ut_2[C_2P(C_2'G_2'R_{12} - (I-H_1G_1C_1-H_2G_2C_2)'\Lambda H_2] = 0.$$

For second order sufficiency conditions, we need

$$\frac{d^2}{d\varepsilon^2} L(G_1 + \varepsilon\Delta G_1, G_2, P, \Lambda)\Big|_{\varepsilon=0} > 0$$

$$\frac{d^2}{d\varepsilon^2} L(G_1, G_2 + \varepsilon\Delta G_2, P, \Lambda)\Big|_{\varepsilon=0} > 0$$

we need

$$\therefore\ tr(C_iPC_i')(\Delta G_i'(R_{1i} + H_i'\Lambda H_i)\Delta G_i) > 0 \qquad i = 1,2.$$

Sufficient conditions are

$$P > 0 \qquad \forall \Delta G_i \in \text{lower block triangular}$$

$$R_{1i} + H_i'\Lambda H_i > 0 \qquad i = 1,2.$$

Q.E.D.

$Y(y_1) = G_1 C_1 X$

END

FILMED

3-83

DTIC